Winter 2021

SARS-CoV-2: An Investigation on Mutagenicity and its Effects on Infectivity and Mortality

Tyler Elliott Silverwood
University of New Hampshire, Durham

SARS-CoV-2: An Investigation on Mutagenicity and its Effects on Infectivity and Mortality

Senior Honors Thesis

Tyler Silverwood

Advisor: Dr. W. Kelley Thomas

December 2021

University of New Hampshire, Durham, NH

College of Life Sciences and Agriculture

Department of Molecular, Cellular, and Biomedical Sciences

Contact: tyleresilverwood@gmail.com or tyler.silverwood@unh.edu

Abstract

SARS-CoV-2, the etiological agent of the COVID-19 pandemic, has rapidly become a worldwide public health concern. Classified as a betacoronavirus, it is the third human coronavirus (HCoV) to emerge in the 21st century that causes severe disease, alongside SARS-CoV and MERS-CoV. The genome consists of open reading frames encoding accessory proteins and four structural proteins, including the spike protein which is a key determinant of host cell tropism. Mutations within the genome, particularly the spike gene, have been linked in-vitro to increased binding affinity to the human receptor angiotensin-converting enzyme 2 (hACE2), increased fitness in human hosts, and immune evasion. Here, both previously studied and novel mutations were correlated with an increase in case and death rates in California, U.S.A. Eleven mutations, occurring both within nonstructural proteins as well as the spike protein, were found to correlate with significantly increased deaths, while one mutation in the open reading frame 3a (ORF3a) gene correlated with both increased cases and death rates.

Introduction

In late December 2019, clusters of pneumonia patients began to emerge in Wuhan, China, centered around wet seafood markets. A novel coronavirus, initially named 2019-nCoV, was determined to be the etiological agent of the disease (Zhu et al., 2020). In March 2020, it was officially designated as SARS-CoV-2 by the International Committee on Taxonomy of Viruses (Zhu et al., 2020). The virus has spread throughout the world, and as of the time of writing, has resulted in over 265 million infections and 5.2 million deaths according to the World Health Organization Coronavirus COVID-19 Dashboard, with a mortality rate of about 2%. It is the seventh coronavirus discovered to be capable of infecting humans (HCoVs) and is classified as a betacoronavirus (Zhu et al., 2020). Coronaviruses, named for their crown – or corona – appearance when viewed under electron microscopy, are single-stranded positive-strand RNA viruses with a typical genome length of around 30kb, making coronavirus genomes the largest of all RNA viruses (Lee et al., 1991; R. & Sonia, 2005). Typical betacoronavirus genomes consist of a 5' untranslated region (UTR), open reading frames encoding nonstructural accessory proteins, spike (S), envelope (E), membrane (M), nucleocapsid (N), and a 3' UTR (Zhu et al., 2020). The spike protein is a key determinant of host tropism (Kang et al., 2021), and is comprised of two subunits. The S1 subunit contains a receptor-binding domain (RBD) which recognizes and binds with the receptor, and S2 mediates membrane fusion (Korber et al., 2020).

HCoVs are ubiquitous in society; four viruses circulate as the common cause of the common cold. The first HCoV was isolated in 1965 from common cold patients (Tyrell & Bynoe, 1965), and one year later was classified as HCoV-229E (Hamre & Procknow, 1966). Soon afterward, HCoV-OC43 was isolated from patients experiencing similar symptoms as HCoV-229E. However, they were found to have no serological cross-reactivity (McIntosh et al.,

1967), suggesting that each virus has distinct surface proteins and therefore unique evolutionary pathways. HCoV-HKU1, which emerged in 2005, has rapidly spread across the globe and similarly produces mild symptoms (Woo et al., 2005). HCoV-NL63 typically induces minor symptoms in immunocompetent patients and can cause severe respiratory illness in young children and immunocompromised individuals (Ebihara et al., 2005). Coronaviruses have been found to proliferate in animals such as bats, which act as reservoirs for the virus. Recombination of viruses is typical in these hosts, giving rise to novel viruses (Su et al., 2016).

In 2002, the world faced a new degree of coronavirus threat, as clusters of deadly pneumonia emerged in China, with the disease later labeled severe acute respiratory syndrome (SARS). A novel HCoV was soon discovered to be the causative agent of the disease and was classified SARS-CoV (Peiris et al., 2003). After plaguing the world for months, the virus vanished, with 8098 recorded probable cases and leaving almost 10% of those patients dead (Stadler et al., 2003). SARS-CoV is believed to have arisen from spillover events of SARS-related (SARSr) coronaviruses that typically circulate in bats. SARSr-Rh-BatCoV, a genetically similar virus, was discovered to circulate in high levels in Chinese horseshoe bats in the region where SARS-CoV emerged. Given the genetic identity, it is likely that horseshoe bats act as the reservoir for SARS-CoV-like viruses, while other species capable of being infected with SARSr-CoVs, such as civets, are the intermediate hosts (Lau et al., 2010). Ten years later, in 2012, another deadly coronavirus emerged: the Middle Eastern respiratory syndrome virus, MERS-CoV, which spreads in small clusters throughout the world, never reaching sustained community transmission (Raj et al., 2014). Compared to other HCoVs, MERS-CoV has an astonishingly deadly; as of 2017, there were 2040 confirmed cases and 712 deaths; a mortality rate of ~35% (Chafekar & Fielding, 2018). The virus's natural reservoir has been discovered to be camels – in

2015, a novel MERS-CoV lineage showing high levels of recombination was traced back to the animal (Su et al., 2016).

Much debate has occurred surrounding the origins of SARS-CoV-2. In January 2020, excess pneumonia deaths were localized around the Wuhan markets, rather than other hypothesized sources such as the Wuhan Institute of Virology. This suggests that the source of the virus was the markets themselves (Holmes et al., 2021). Like SARS-CoV, SARS-CoV-2 shows a high degree of nucleotide identity with bat viruses. Bat CoV RaTG13, which was isolated in 2013, shares 96% identity, including a 97% amino acid identity of the spike protein (Kang et al., 2021). SARSr-CoVs bat-SL-CoVZC45 and bat-SL-CoVZCX21, isolated in 2018, share an 88% identity (Lu et al., 2020). However, due to the geographical distance between the Wuhan markets and the source of these isolates, it is probable that like SARS-CoV, SARS-CoV-2 had an intermediate host that increased fitness for human hosts. Pangolins, which were present at the markets, have been proposed as that intermediate host: pangolin SARSr-CoVs have been found that have a ~97.4% amino acid identity of the spike protein (Kang et al., 2021). SARS-CoV-2, SARS-CoV, and MERS-CoV likely share a common ancestor, as the newer virus shares a 79% identity to SARS-CoV and a 50% identity to MERS-CoV (Lu et al., 2020).

Nucleotide mutational rate is relatively consistent across the three deadly HCoVs. Across all CoVs, there is an average mutational rate of ~10^{-4} per year per site (Pyrc et al., 2006). SARS-CoV's mutational rate lies between 0.80-2.38x10^{-3} per year per site (Zhao et al., 2004), while MERS-CoV's rate is estimated to be 1.12x10^{-3} per year per site (Cotten et al., 2014), and SARS-CoV-2's is estimated to be ~1.8x10^{-3} per year per site (Li et al., 2020). The SARS-CoV-2 spike protein has a higher rate of mutation when compared to the rest of the genome. On a per replication basis, the entire genome has a rate of 1.25x10^{-6} mutations per site, while the S region

has a rate one magnitude higher of $1.6x10^{-5}$ mutations per site per replication (Amicone et al., 2021). The increased mutational rate of the S region likely confers an advantage to SARS-CoV-2 as it adjusts to human cells. Early selective sweeps within the S region of the genome have been hypothesized to have contributed to the zoonotic leap from the intermediate host to humans, including mutations T372A (Kang et al., 2021) and D614G (Holmes et al., 2021) that were both found to vastly increase human cellular binding efficiency. Despite key mutations, both SARS-CoV-2 and SARS-CoV have similar spike RBDs, and utilize human angiotensin-converting enzyme 2 (hACE2) as their cellular receptors (Lu et al., 2020).

Given SARS-CoV-2's novelty to humans, it is expected that the virus will change as it adapts to its new host. Improvements to viral fitness will ultimately present themselves as changes in infectivity and mortality, measured via case and death rates. In this study, isolates from California, U.S.A. will be analyzed to determine mutations that may affect viral fitness. Mutations that occur within key regions of the SARS-CoV-2 genome that affect cell entry and replication, such as the spike, are hypothesized to increase both case and death rates.

Methodology

Genomes isolated from samples in California, U.S.A. were downloaded from the National Library of Medicine National Center for Biotechnology Information (NCBI) SARS-CoV-2 resources hub for collection dates between January 2020 and February 2021. Sample density ranged from 7 (January 2020) to 3600 (January 2021). Monthly case and death rates were downloaded from the COVID-19 Data Repository by the Center of Systems Science and Engineering at Johns Hopkins University, and extracted utilizing a homebrew python script

extractData.py. Wtdbg2 assembler was utilized to create consensus sequences for each month. Samples were then assembled into a master consensus sequence to which each month's consensus sequence was aligned, utilizing MAFFT version 7 (Katoh & Standley, 2013). CodonCode Aligner version 9.0.2 (CodonCode, Dedham, MA) was utilized to parse samples for mutations, which were then exported to Excel. Genotypes with a mutation that appears in more than two months' consensus sequences were analyzed. Workflow performed again, on a weekly instead of a monthly basis. Genomes isolated from California, U.S.A. were downloaded from the NCBI from the week beginning 04/01/2020 to 03/30/2021, producing 52 weekly consensus sequences. Python script extractDataWeekly.py utilized to extract weekly case and death data. Final sequences were aligned with the publicly available reference genome MN90847 published by Ensembl on the European Nucleotide Archive (EMBL-EBI, Hinxton, Cambridgeshire, UK) utilizing EMBOSS Stretcher (Myers & Miller, 1988). To analyze structural differences between selected mutants, Phyre^2 was utilized for .pbd generation and structure predictions (Kelley et al., 2015). Molecular graphics and analyses were performed using UCSF Chimera from the Resource for Biocomputing, Visualization, and Informatics and the University of California, San Francisco (Pettersen et al., 2004).

Results and Discussion

Monthly Analysis

One point mutation was found to correlate with significantly increased case and death rates, while another point mutation was found to correlate with increased death, but not case, rates. G25563T correlated with increased case rates ($p=0.011$) and death rates ($p=0.018$); while

T27890G correlated with increased death rates (p=0.031) but not case rates (p=0.31). Other point mutations that were present in multiple month's consensus sequences but found to not significantly correlate with neither case nor death rates were T1054C and G20263A. Complete data can be found in *California Data.xlsx*.

Weekly Analysis

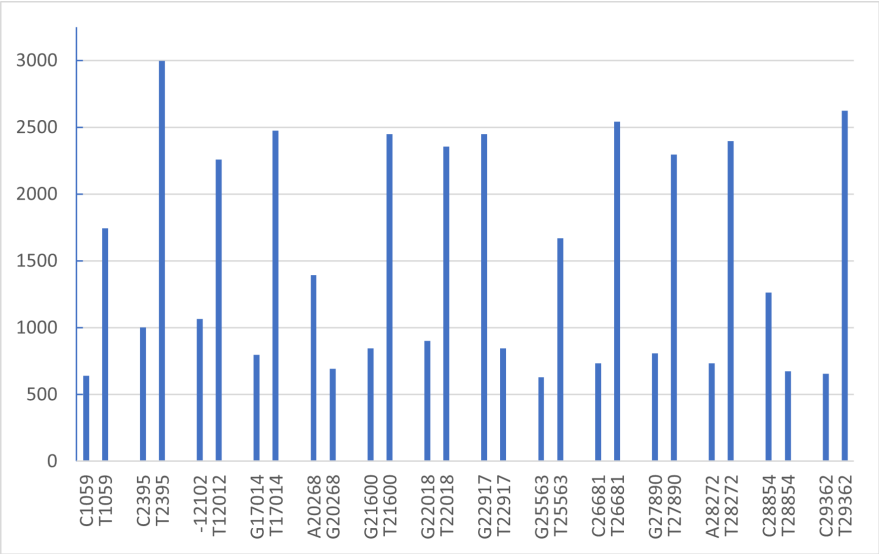

Figure 1. Selected genomic isoforms and their average weekly deaths

One point mutation, G25563T (p=0.022) was found to correlate with significantly increased case rates, while 10 mutations correlated with significantly increased death rates: C1059T (p=0.0002), C2395T (p=0.005), thymine insertion at 12102 (p=0.0388), G17014T (p=0.00004), G21600T (p=0.0001), G22018T (p=0.0007), T22917G (p=0.0001), G25563 (p=0.000186), C26681T (p=0.000007), C29362T (p=0.000002). Three mutants were discovered that appeared in fall 2020-early winter 2021, that correlated with less deaths than the wild-type that reappeared later: A20268G (p=0.0017), A28272T (p=0.00008), C28854T (p=0.0021).

Around 150 other mutations were found to not significantly correlate with increased case or death rates. Complete data can be found in California Correlation Weekly.xlsx.

Genome Mutation	Amino Acid Mutation	Protein	Predicted Structure Change(s)
C1059T	T84I	Nsp2	β-sheets change to α-helices around residues 100 and 120.
C2395T	Synonymous mutant	Nsp2	None
12102T (insertion)	AIA > L-P	Nsp8	Loss of N-terminal α-helix
G17014T	D10Y	Nsp13 helicase domain	C-terminal shift of α-helix around residue 10; shortening of α-helix around residue 110
G21600T	S13I	S	Inward folding of β-sheets around residue 680; loss of α-helix around residue 250
G22018T	W152C	S	Loss of α-helix around residue 250
T22917G	L452R	S	Hydrophobicity change in RBD
G25563T	Q57H	ORF3a	Hydrophobicity change around H57
C26681T	Synonymous mutant	M	None
T27890G	n/a	n/a	n/a (Occurs in noncoding region)
C29362T	Synonymous mutant	N	None

Table 1. Selected mutations and their effect on SARS-CoV-2 proteins.

Stabilizing mutation of ORF3a may convey increased pathogenicity to SARS-CoV-2

Single nucleotide polymorphism G25563T was correlated with significantly increased cases and deaths. This nucleotide occurs in position 171 of the open reading frame 3a (ORF3a) segment of the genome, resulting in a nonsynonymous mutation of amino acid 57 from glutamine to histidine (Q57H). ORF3a encodes a 275 amino acid, six-subunit protein consisting of three transmembrane domains and an intracellular cytoplasmic domain, localized to the region of the cell containing the Golgi apparatus (Padhan et al., 2007; Kern et al., 2021). ORF3a's function and pathogenicity have been widely studied, primarily through its SARS-CoV analog. However, its precise mechanism of action is unknown. One transmembrane domain contains a potassium-sensitive ion channel that is believed to promote virus release, and inhibition of this

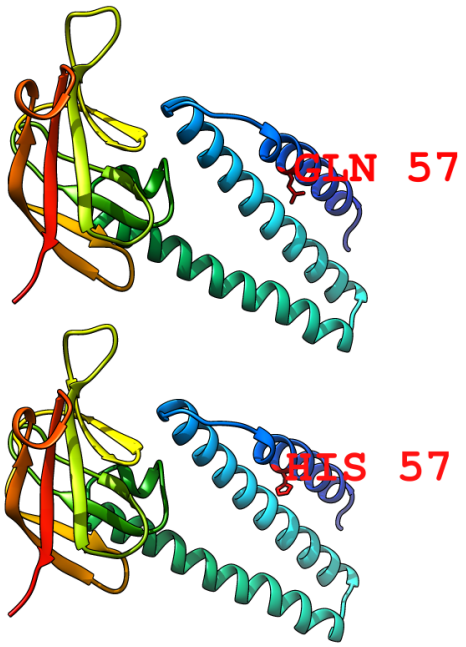

Figure 2. ORF3a protein, highlighting residue 57 with molecule depiction (wild-type, top; mutant, bottom)

domain using siRNA lead to significantly decreased viral load (Lu et al., 2006). The intracellular domain of ORF3a has been shown to contain binding sites for caveolin proteins, which contribute to the formation of caveolae, membrane microdomains that are responsible for the intake of small molecules and play a role in viral entry. Specifically, caveolin-1 has been shown to regulate extracellularly regulated kinase (ERK) and inducible nitric oxide synthase (iNOS), which are involved in the cellular response to viral infection (Padhan et al., 2007). Through this second mechanism, ORF3a inhibiting caveolin-1 action may interfere with cellular defence processes. ORF3a has been implicated in pro-apoptotic pathways, inducing apoptosis through cleavage and activation of caspase-8. Interestingly, this mechanism in SARS-CoV-2 is significantly weaker than in SARS-CoV (Ren et al., 2020). This suggests that the mutations in the ORF3a are responsible for the decreased mortality of the newer virus when compared to its relative, and may play a role in the high rate of asymptomatic infection. Along with proapoptotic activity, ORF3a promotes a cytokine response to infection through activation of inflammatory mediators C-Jun N-terminal kinase (JNK), nuclear factor kappa B (NF-κB), and interleukin 8 (IL-8) promoters (Kanzawa et al., 2006). It additionally has been shown to activate pro-IL-1β gene transcription (Siu et al.,

2019). The pro-inflammatory action of ORF3a likely contributes to hyper-inflammation and cytokine storm, a key cause of COVID-19 mortality (Mehta et al., 2020).

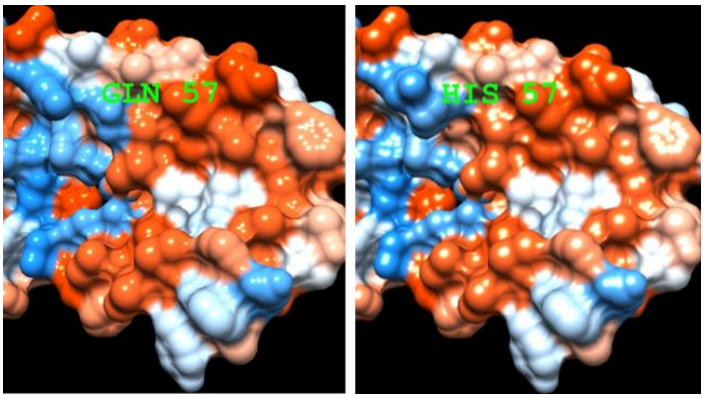

Figure 3. Hydrophobicity surface of SARS-CoV-2-ORF3a, zoomed on residue 57 (wild-type, left; Q57H, right. Blue, hydrophilic; red, hydrophobic)

Amino acid 57 occurs in the center of the ORF3a protein, near the central pore and lower tunnel segments. In wild-type SARS-CoV-2, amino acid 57 occurs as the uncharged polar amino acid glutamine, while the mutation to histidine causes this residue to become charged. This ultimately results in the surface becoming more hydrophobic with a minor change to surface geometry. Two separate studies, by Hassan et al. (2021) and Bianchi et al. (2021) have utilized simulation software Meta-SNP and DynaMut, respectively, to show that this mutation conveys a stabilizing effect to the protein and likely increases pathogenicity. Stabilization of ORF3a may potentiate the pathogenic effects outlined above, ultimately leading to the observed increase in infections and deaths where the Q57H mutation is dominant.

Nonstructural protein 2 (nsp2) may promote a pro-viral cell environment

Mutation of cytosine to thymine at base 1059 within the open reading frame 1ab (ORF1ab) region correlates with an increased rate of death. ORF1ab is a polyprotein sequence consisting of overlapping reading frames *a* and *ab*, encoding 16 nonstructural proteins. C1059T is a nonsynonymous mutation occurring in nonstructural protein 2 (nsp2), resulting in the

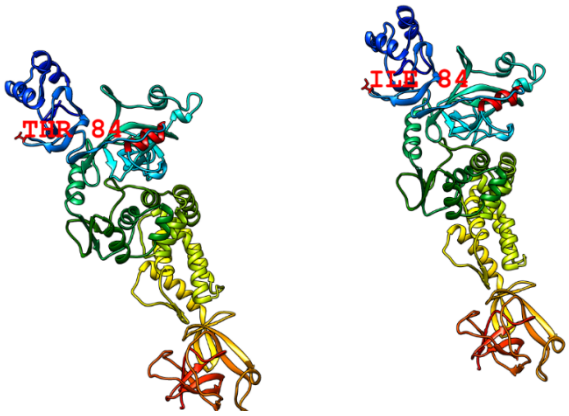

Figure 4. Nonstructural Protein 2, highlighting residue 84 with molecule depiction (wild-type, left; mutant, right)

conversion of the polar uncharged amino acid threonine to the nonpolar uncharged residue serine (T84I).

Nsp2 has a wide range of hypothesized effects, ranging from affecting mitochondria biogenesis to inhibiting apoptosis, and its sequence varies greatly between HCoVs (Davies et al., 2020). The SARS-CoV nsp2 analog has been found to form complexes with at least 11 host proteins in HEK293 cells. The strongest interactions are with cellular signaling proteins prohibitin 1 and 2 (PHB1/PHB2), which localize to both the mitochondrial inner membrane and nucleus (Tatsuta et al., 2005; T. et al., 2009). PHB1 and PHB2 are believed to play a role in the cell cycle shift from G1 to S phase, conveying anti-proliferative effects. Their mechanism of action appears to be in the repression of E2F, a family of transcription factors that consists of both cell cycle progression activators and repressors (Wang et al., 1999). Overexpression of activator factors in the E2F family has been shown to cause cells to reenter S phase despite the presence of other signals for cycle arrest (Dimova & Dyson, 2005). Therefore, if the interaction of nsp2 and prohibitins is repressive, E2F repression may be reduced. While this would not necessarily result in overexpression of activator E2F factors, reduced repression may result in similar effects including the upregulation of S phase cell machinery, an optimal outcome for SARS-CoV-2 which relies on this machinery for its replication. Nsp2 may similarly promote a welcoming cell environment through interactions with stomatin-like protein 2 (STOML2), a mitochondrial regulator. Increased STOML2 levels within a cell have been linked to increased ATP production and inhibition of apoptosis pathways, both beneficial for viral replication (A. et

al., 2011). Nsp2 may potentiate these effects of STOML2, further promoting a pro-viral environment.

Nonstructural protein 8 (nsp8) enhances RNA polymerase activity

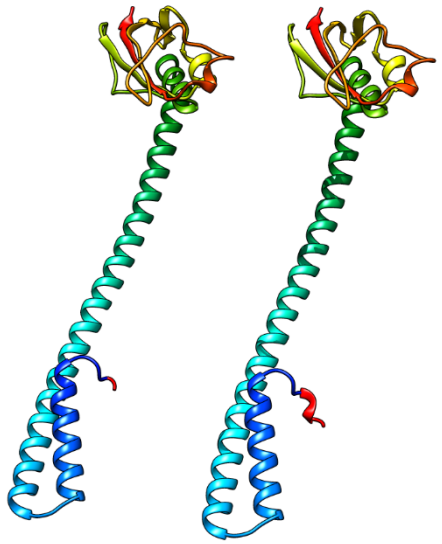

Figure 5. Nonstructural Protein 8 (mutant, left; wild-type, right. Difference highlighted in red)

Thymine insertion at position 12102 occurs in the ORF1ab segment of the genome, at base pair 11 of the segment encoding nonstructural protein 8 (nsp8), a golf club-shaped accessory protein. This results in the first three codons translating to N-STOP-P instead of AIA, shortening the polypeptide by two residues. Nsp8 appears to assist in viral genome replication, as it was found to form a complex with nsp7 and nsp12, a catalytic accessory protein containing an RNA-dependent RNA polymerase domain (RdRp) (Ahn et al., 2012). This complex is responsible for replicating the viral genome within the host cell, and the subunits are relatively well-conserved between SARS-CoV-2 and SARS-CoV. Two copies of Nsp8 form "sliding poles" along the RNA's path of egress from the complex, assisting in preventing the dissociation of the newly formed genome from the polymerase (Hillen et al., 2020). This segment has an internal diameter of about 30Å, with the N-terminal domain (NTD) – consisting of two α-helices – on the outer edge of the complex (Zhai et al., 2005). Both nsp7 and nsp8 are essential to nsp12's RdRp function, enhancing its activity (Subissi et al., 2014). While the entire complex is relatively highly conserved between the two HCoVs, SARS-CoV-2's polymerase complex is significantly less efficient and stable than its SARS-CoV relative, operating at about a 35% relative efficiency (Peng et al., 2020). There are myriad reasons behind this difference, including

that the machinery is likely optimized for the virus's natural reservoir and functions just sufficiently enough for human pathogenicity; as the virus evolves, one could expect the complex's efficiency and stability to increase. Interestingly, substituting SARS-CoV's nsp8 analog into the complex boosted its efficiency by 2.1 times, highlighting this accessory protein's importance in the infection process (Peng et al., 2020). Nsp8 has also been found to interact with the ORF6 protein in SARS-CoV, and while their exact interaction is unknown, ORF6 itself has been implicated in coronavirus lethality (Pewe et al., 2005). Given this mutation's correlation with increased death rates, it is likely that the subsequent structural change in nsp8 increases the polymerase's efficiency and may potentiate ORF6's lethal effects.

Nonstructural protein 13 (nsp13) interacts with the SARS-CoV-2 RdRp Holoenzyme

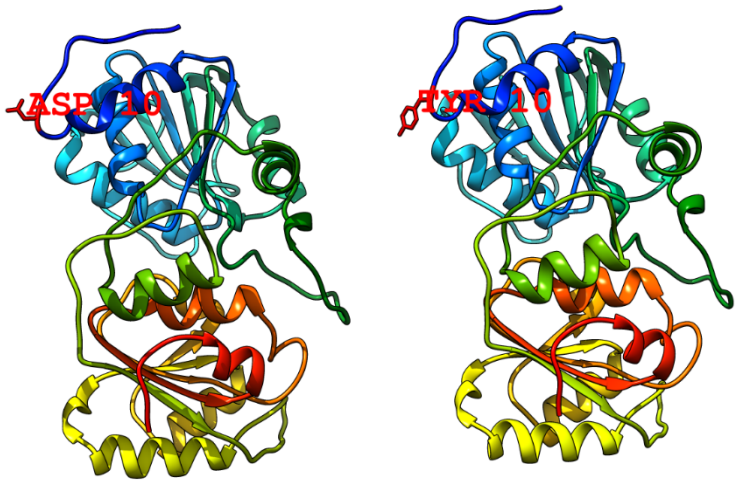

Figure 6. Nonstructural Protein 13 Helicase Domain, highlighting residue 10 with molecule depiction (wild-type, left; mutant, right)

SARS-CoV-2 nonstructural protein 13 (nsp13) is a ~67kDa accessory protein encoded within the ORF1ab polyprotein sequence (Tanner et al., 2003). It consists of four distinct domains: an N-terminal zinc-binding domain (ZBD), a helical "stalk" domain, a β-barrel, and a helicase domain (Newman et al., 2021). Mutation of G to T at nucleotide 17014 has been correlated with higher death rates and results in the transformation of helicase residue D10 to Y (D10Y). In SARS-CoV, whose nsp13 analog differs by only one amino acid (Newman et al., 2021), nsp13 functions as a helicase, unwinding approximately 280bp per second; it was found to interact with

the RdRp domain of nsp12, which doubles the unwinding rate (Adedeji et al., 2012; Jia et al., 2019). This helicase action occurs exclusively in the 5' to 3' direction (Shu et al., 2020). Nsp13 additionally has nucleotide triphosphate (NTP)-ase activity, showing a preference for ATP and GTP (Shu et al., 2020). Nsp13 interacts strongly with the entire SARS-CoV-2 replication-transcription complex (RTC) of nsp7, nsp8, and nsp12 (RdRp holoenzyme), forming the complex nsp13-RTC. While nsp13 helicase translocation occurs in the 5'-3' direction, RdRp translocation occurs 3'-5'. This indicates that nsp13's primary role is in a process referred to as backtracking, where the RTC slides backward on the nucleic acid (Chen et al., 2020). Backtracking has numerous functions, notably including maintaining genomic fidelity through proofreading and the processing of sub-genomic RNA transcripts (Nudler, 2012). Proofreading in SARS-CoV-2 involves the nsp10-nsp14 complex, which contains an N-terminal exonuclease (ExoN) domain that is normally unable to access the nucleotide buried in the RdRp active site (Subissi et al., 2014). Backtracking causes the nsp13 helicase to translocate the RNA strand backward, resulting in the erroneous single-stranded 3' RNA to be moved out of the RdRp active site and into the NTPase active site, allowing it to be accessed by other enzymes such as the nsp10-nsp14 complex (Malone et al., 2021). Backtracking additionally confers protection against mutagenic antivirals, such as remdesivir, and has been widely studied in HCoVs (Smith et al., 2013). This process evolved independently in both RdRp and DNA-dependent RNA polymerase (DdRp), highlighting its importance in replication and transcription (Malone et al., 2021).

Nsp13 additionally has been found to promote a pro-viral environment through the inhibition of interferon and NF-κB. It has been found to associate with TBK1, a regulatory kinase that is involved in IRF3 and NF-κB signaling pathways; phosphorylated TBK1 and IRF3 were found to exist in lower levels in cells with nsp13 overexpression (Vazquez et al., 2021).

Additionally, it was found to inhibit IFN-β promoter activity (Lei et al., 2020). These effects appear to act counter to the pro-inflammatory action of the ORF3a protein discussed above. Given nsp13's key role in replication fidelity, transcription, and cell defense mechanism disruption, it is being actively investigated as an antiviral target.

Spike (S) is a Key Determinant of Cellular Tropism, Infectivity, and Immune Response

The S protein of CoVs is perhaps the most widely studied of all CoV proteins, rightfully so. S, colloquially known as the spike protein, is a ~200kDa protein consisting of an extracellular N-terminal end, a transmembrane domain, and an intracellular C-terminal end (Bosch et al., 2003). S exists as homotrimers on the surface of the SARS-CoV-2 virion and is a viral fusion protein that mediates host cell attachment and fusion of the cell-virion membranes (Huang et al., 2020) (Jan et al., 2003). The S protein contains a receptor-binding domain (RBD), which is a variable region that is the key determinant of receptor tropism. Within the RBD of SARS-CoV-2 is the receptor-binding motif (RBM), a short string of residues that are directly responsible for binding to the host cell's receptors (Li et al., 2005). Both SARS-CoV and SARS-CoV-2 utilize the human angiotensin-converting enzyme 2 (hACE2) as their receptors

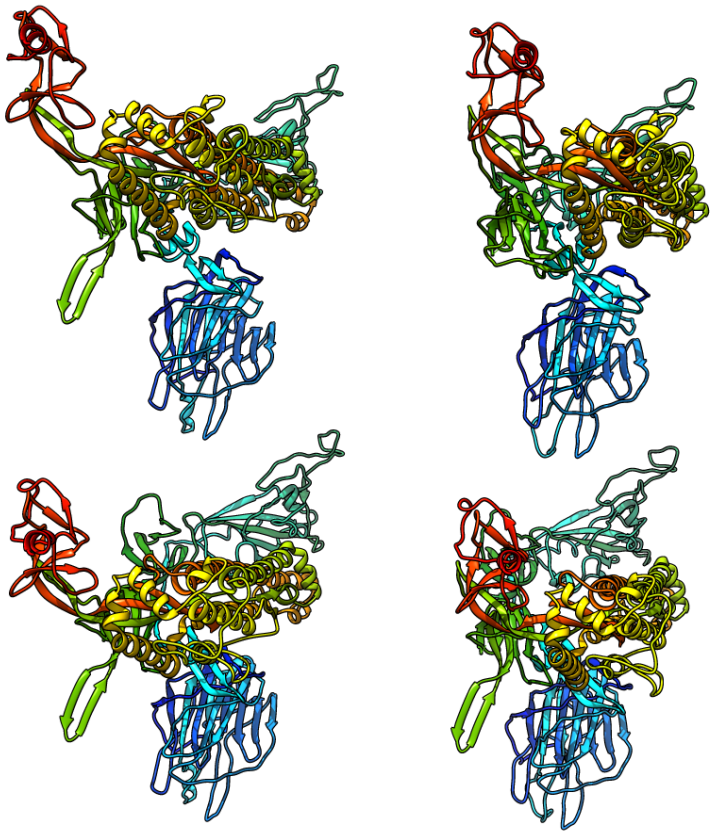

Figure 7. Wild-type SARS-CoV-2 Spike (S) protein and mutants (In clockwise order: wild-type, S13I, W152, L452R)

and pathway into a cell (Fang et al., 2005; Wuze et al., 2008) – consistent with the fact that the two viruses share a high sequence identity of their RBM (Hoffmann et al., 2020). However, SARS-CoV-2 has a 10-to-20-fold higher affinity to hACE2 compared to SARS-CoV, suggesting that factors other than the RBM also play a role in cell infection (Daniel et al., 2020). Once S has entered the cell through the hACE2 receptor, a protease located on the host cell membrane, TM protease serine 2 (TMPRSS2), cleaves S into the sub proteins S1 and S2. S1 is responsible for facilitating virion attachment to the host's surface, while S2 carries out membrane fusion (Fehr & Perlman, 2015). A cleavage site at the S1/S2 boundary assists in the processing of S; SARS-CoV-2 contains a Furin-like cleavage site (FCS), which is unique when compared to other HCoVs. The FCS is the result of mutations in the S gene that produce a protein containing the sequence 682-RRAR-685 and appears to dramatically increase SARS-CoV-2's cell infectivity outside of the human respiratory tract (Xia et al., 2020). Given the ubiquity of hACE2 throughout the body (Hamming et al., 2004), the evolution of the FCS likely played a large role in the improvement of SARS-CoV-2's binding affinity and contributes to COVID-19's systemic symptomatology.

Both the S homotrimer and the presence of glycosylation sites give the virus its characteristic crown visual appearance. S contains between 66-87 N-linked glycosylation sites (Watanabe et al., 2020), and the configuration of these sites contributes to host cell tropism, protein folding, and stability (Watanabe et al., 2019). Additionally, glycosylation sites aid in facilitating immune system evasion, by shielding key epitopes from antibodies (Watanabe et al., 2020).

Three nonsynonymous mutations within the S gene have been correlated with increased deaths. Residue 13 occurs within the N-terminal signal region, which plays a role in S localization within the cell (Huang et al., 2020). Mutation S13I has significant effects on the structure of the spike, resulting in an inward folding of the β-sheets around residue 680, as well as the loss of an α-helix around residue 110. This results in a decrease in protein surface area, from the wild-type's $5.3x10^4$ $Å^2$ to $4.96 x10^4$ $Å^2$. Previous literature has linked the S13I mutation, present in the epsilon variants B.1.427 and B.1.429, with a significant decrease in antibody binding affinity to S. This is believed to occur due to S13I inducing a shift in the NTD antigenic site (Matthew et al., 2021); the immune evasion induced by this shift may be responsible for the noted correlation with increased deaths. W152 occurs within the NTD, and results in the loss of the circa-residue 110 α-helix. The NTD appears to assist in the association between the RBD and hACE2, and is a novel target of antiviral drugs (Hamre & Procknow, 1966); mutations in this region may strengthen the binding between the virion and its host cell receptor. L452R occurs within the critical RBM domain, changing the non-polar and hydrophobic leucine to the polar hydrophilic amino acid arginine. This specific mutation is present in a subset of Alpha variant (lineage B.1.1.7) isolates, which became a dominant variant during the 2020 winter surge (Washington et al., 2021). Our findings of a correlation between

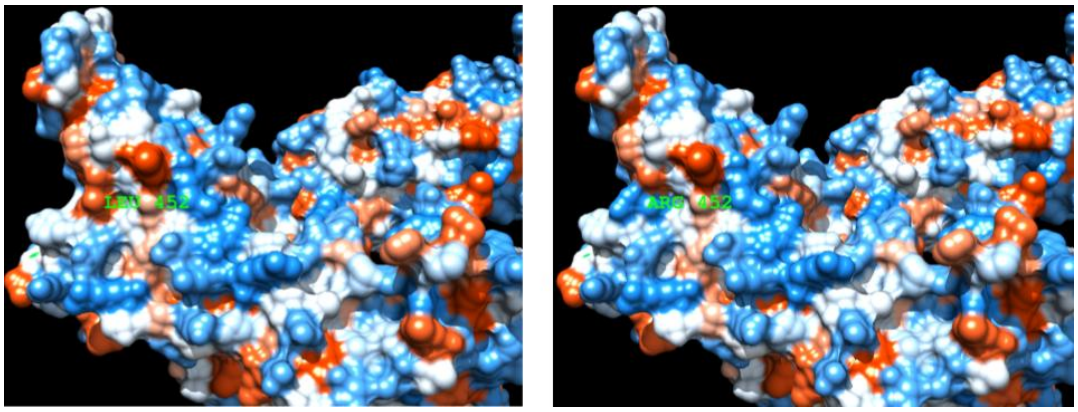

Figure 8. Hydrophobicity surface of SARS-CoV-2-S, zoomed on the RBD (wild-type, left; L452R, right. Blue, hydrophilic; red, hydrophobic)

L452 and increased death rates also support previous findings that the mutation improves S-hACE2 affinity and increases S stability and replication (Motozono et al., 2021).

As three deadly HCoVs have emerged within the last two decades, it is important to understand every aspect of the viruses' mechanisms of action, evolution, and mutational behavior. Given the ubiquity of SARS-related viruses circulating within animal reservoirs, it is simply a matter of time before another spillover event occurs and a new HCoV emerges, whether it causes deadly diseases like SARS and MERS, or be as relatively benign as the common cold. Increased surveillance of known reservoirs and intermediate hosts may assist in preventing the next pandemic HCoV, and a complete understanding of how mutations may affect infectivity and mortality is necessary for the rapid identification of concerning isolates. The ongoing COVID-19 pandemic has also given rise to variants of concern harboring multiple mutations, such as the recently discovered omicron variant that harbors many mutations in the S gene RBD, as well as mutations that may impact the FCS (Karim & Karim, 2021). Robust surveillance programs screening of mutations of concern, such as the mutations discussed in this paper, will aid in detection and preventing the worldwide spread of more variants of concern.

References

A., C. D., D., L. C., M., E. I., A., C. L., G., K. M., Bo, L., H., B. E., D., D. S., M., H. G., & Joaquín, M. (2011). Stomatin-Like Protein 2 Binds Cardiolipin and Regulates Mitochondrial Biogenesis and Function . *Molecular and Cellular Biology*, *31*(18), 3845–3856. https://doi.org/10.1128/MCB.05393-11

Adedeji, A. O., Marchand, B., te Velthuis, A. J. W., Snijder, E. J., Weiss, S., Eoff, R. L., Singh, K., & Sarafianos, S. G. (2012). Mechanism of Nucleic Acid Unwinding by SARS-CoV Helicase. *PLOS ONE*, *7*(5), e36521. https://doi.org/10.1371/journal.pone.0036521

Ahn, D.-G., Choi, J.-K., Taylor, D. R., & Oh, J.-W. (2012). Biochemical characterization of a recombinant SARS coronavirus nsp12 RNA-dependent RNA polymerase capable of copying viral RNA templates. *Archives of Virology*, *157*(11), 2095–2104. https://doi.org/10.1007/s00705-012-1404-x

Amicone, M., Borges, V., Alves, M. J., Isidro, J., Zé-Zé, L., Duarte, S., Vieira, L., Guiomar, R., Gomes, J. P., & Gordo, I. (2021). Mutation rate of SARS-CoV-2 and emergence of mutators during experimental evolution. *BioRxiv*, 2021.05.19.444774. https://doi.org/10.1101/2021.05.19.444774

Bianchi, M., Borsetti, A., Ciccozzi, M., & Pascarella, S. (2021). SARS-Cov-2 ORF3a: Mutability and function. *International Journal of Biological Macromolecules*, *170*, 820–826. https://doi.org/10.1016/j.ijbiomac.2020.12.142

Bosch, B. J., van der Zee, R., de Haan, C. A. M., & Rottier, P. J. M. (2003). The coronavirus spike protein is a class I virus fusion protein: structural and functional characterization of the fusion core complex. *Journal of Virology*, *77*(16), 8801–8811. https://doi.org/10.1128/jvi.77.16.8801-8811.2003

Chafekar, A., & Fielding, B. C. (2018). MERS-CoV: Understanding the Latest Human Coronavirus Threat. *Viruses*, *10*(2), 93. https://doi.org/10.3390/v10020093

Chen, J., Malone, B., Llewellyn, E., Grasso, M., Shelton, P. M. M., Olinares, P. D. B., Maruthi, K., Eng, E. T., Vatandaslar, H., Chait, B. T., Kapoor, T. M., Darst, S. A., & Campbell, E. A. (2020). Structural Basis for Helicase-Polymerase Coupling in the SARS-CoV-2 Replication-Transcription Complex. *Cell*, *182*(6), 1560-1573.e13. https://doi.org/10.1016/j.cell.2020.07.033

Cotten, M., Watson, S. J., Zumla, A. I., Makhdoom, H. Q., Palser, A. L., Ong, S. H., Al Rabeeah, A. A., Alhakeem, R. F., Assiri, A., Al-Tawfiq, J. A., Albarrak, A., Barry, M., Shibl, A., Alrabiah, F. A., Hajjar, S., Balkhy, H. H., Flemban, H., Rambaut, A., Kellam, P., & Memish, Z. A. (2014). Spread, circulation, and evolution of the Middle East respiratory syndrome coronavirus. *MBio*, *5*(1). https://doi.org/10.1128/mBio.01062-13

Daniel, W., Nianshuang, W., S., C. K., A., G. J., Ching-Lin, H., Olubukola, A., S., G. B., & S., M. J. (2020). Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science*, *367*(6483), 1260–1263. https://doi.org/10.1126/science.abb2507

Davies, J. P., Almasy, K. M., McDonald, E. F., & Plate, L. (2020). Comparative Multiplexed Interactomics of SARS-CoV-2 and Homologous Coronavirus Nonstructural Proteins Identifies Unique and Shared Host-Cell Dependencies. *ACS Infectious Diseases*, *6*(12), 3174–3189. https://doi.org/10.1021/acsinfecdis.0c00500

Dimova, D. K., & Dyson, N. J. (2005). The E2F transcriptional network: old acquaintances with new faces. *Oncogene*, *24*(17), 2810–2826. https://doi.org/10.1038/sj.onc.1208612

Dong E, Du H, Gardner L. An interactive web-based dashboard to track COVID-19 in real time. Lancet Inf Dis. 20(5):533-534. doi: 10.1016/S1473-3099(20)30120-1

Ebihara, T., Endo, R., Ma, X., Ishiguro, N., & Kikuta, H. (2005). Detection of human coronavirus NL63 in young children with bronchiolitis. *Journal of Medical Virology*, *75*(3), 463–465. https://doi.org/10.1002/jmv.20289

Fang, L., Wenhui, L., Michael, F., & C., H. S. (2005). Structure of SARS Coronavirus Spike Receptor-Binding Domain Complexed with Receptor. *Science*, *309*(5742), 1864–1868. https://doi.org/10.1126/science.1116480

Fehr, A. R., & Perlman, S. (2015). Coronaviruses: an overview of their replication and pathogenesis. *Methods in Molecular Biology (Clifton, N.J.)*, *1282*, 1–23. https://doi.org/10.1007/978-1-4939-2438-7_1

Hamming, I., Timens, W., Bulthuis, M. L. C., Lely, A. T., Navis, G. J., & van Goor, H. (2004). Tissue distribution of ACE2 protein, the functional receptor for SARS coronavirus. A first step in understanding SARS pathogenesis. *The Journal of Pathology*, *203*(2), 631–637. https://doi.org/10.1002/path.1570

Hamre, D., & Procknow, J. J. (1966). A new virus isolated from the human respiratory tract. *Proceedings of the Society for Experimental Biology and Medicine. Society for Experimental Biology and Medicine (New York, N.Y.)*, *121*(1), 190–193. https://doi.org/10.3181/00379727-121-30734

Hassan, S. S., Attrish, D., Ghosh, S., Choudhury, P. P., & Roy, B. (2021). Pathogenic perspective of missense mutations of ORF3a protein of SARS-CoV-2. *Virus Research*, *300*, 198441. https://doi.org/10.1016/j.virusres.2021.198441

Hillen, H. S., Kokic, G., Farnung, L., Dienemann, C., Tegunov, D., & Cramer, P. (2020). Structure of replicating SARS-CoV-2 polymerase. *Nature*, *584*(7819), 154–156. https://doi.org/10.1038/s41586-020-2368-8

Hoffmann, M., Kleine-Weber, H., Schroeder, S., Krüger, N., Herrler, T., Erichsen, S., Schiergens, T. S., Herrler, G., Wu, N.-H., Nitsche, A., Müller, M. A., Drosten, C., & Pöhlmann, S. (2020). SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. *Cell*, *181*(2), 271-280.e8. https://doi.org/10.1016/j.cell.2020.02.052

Holmes, E. C., Goldstein, S. A., Rasmussen, A. L., Robertson, D. L., Crits-Christoph, A., Wertheim, J. O., Anthony, S. J., Barclay, W. S., Boni, M. F., Doherty, P. C., Farrar, J., Geoghegan, J. L., Jiang, X., Leibowitz, J. L., Neil, S. J. D., Skern, T., Weiss, S. R., Worobey, M., Andersen, K. G., … Rambaut, A. (2021). The origins of SARS-CoV-2: A critical review. *Cell*, *184*(19), 4848–4856. https://doi.org/https://doi.org/10.1016/j.cell.2021.08.017

Howe, K. L., Achuthan, P., Allen, J., Allen, J., Alvarez-Jarreta, J., Amode, M. R., Armean, I. M., Azov, A. G., Bennett, R., Bhai, J., Billis, K., Boddu, S., Charkhchi, M., Cummins, C., Da Rin Fioretto, L., Davidson, C., Dodiya, K., El Houdaigui, B., Fatima, R., … Flicek, P. (2021). Ensembl 2021. *Nucleic Acids Research*, *49*(D1), D884–D891. https://doi.org/10.1093/nar/gkaa942

Huang, Y., Yang, C., Xu, X., Xu, W., & Liu, S. (2020). Structural and functional properties of SARS-CoV-2 spike protein: potential antivirus drug development for COVID-19. *Acta Pharmacologica Sinica*, *41*(9), 1141–1149. https://doi.org/10.1038/s41401-020-0485-4

Jan, B. B., Ruurd, van der Z., M., de H. C. A., & M., R. P. J. (2003). The Coronavirus Spike Protein Is a Class I Virus Fusion Protein: Structural and Functional Characterization of the Fusion Core Complex. *Journal of Virology*, *77*(16), 8801–8811. https://doi.org/10.1128/JVI.77.16.8801-8811.2003

Jia, Z., Yan, L., Ren, Z., Wu, L., Wang, J., Guo, J., Zheng, L., Ming, Z., Zhang, L., Lou, Z., & Rao, Z. (2019). Delicate structural coordination of the Severe Acute Respiratory Syndrome coronavirus Nsp13 upon ATP hydrolysis. *Nucleic Acids Research*, *47*(12), 6538–6550. https://doi.org/10.1093/nar/gkz409

Kang, L., He, G., Sharp, A. K., Wang, X., Brown, A. M., Michalak, P., & Weger-Lucarelli, J. (2021). A selective sweep in the Spike gene has driven SARS-CoV-2 human adaptation. *Cell*, *184*(17), 4392-4400.e4. https://doi.org/10.1016/j.cell.2021.07.007

Kanzawa, N., Nishigaki, K., Hayashi, T., Ishii, Y., Furukawa, S., Niiro, A., Yasui, F., Kohara, M., Morita, K., Matsushima, K., Le, M. Q., Masuda, T., & Kannagi, M. (2006). Augmentation of chemokine production by severe acute respiratory syndrome coronavirus 3a/X1 and 7a/X4 proteins through NF-kappaB activation. *FEBS Letters*, *580*(30), 6807–6812. https://doi.org/10.1016/j.febslet.2006.11.046

Karim, S. S. A., & Karim, Q. A. (2021). Omicron SARS-CoV-2 variant: a new chapter in the COVID-19 pandemic. *The Lancet*, *398*(10317), 2126–2128. https://doi.org/10.1016/S0140-6736(21)02758-6

Katoh, K., & Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution*, *30*(4), 772–780. https://doi.org/10.1093/molbev/mst010

Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N., & Sternberg, M. J. E. (2015). The Phyre2 web portal for protein modeling, prediction and analysis. *Nature Protocols*, *10*(6), 845–858. https://doi.org/10.1038/nprot.2015.053

Kern, D. M., Sorum, B., Mali, S. S., Hoel, C. M., Sridharan, S., Remis, J. P., Toso, D. B., Kotecha, A., Bautista, D. M., & Brohawn, S. G. (2021). Cryo-EM structure of SARS-CoV-2 ORF3a in lipid nanodiscs. *Nature Structural and Molecular Biology*, *28*(7), 573–582. https://doi.org/10.1038/s41594-021-00619-0

Korber, B., Fischer, W. M., Gnanakaran, S., Yoon, H., Theiler, J., Abfalterer, W., Hengartner, N., Giorgi, E. E., Bhattacharya, T., Foley, B., Hastie, K. M., Parker, M. D., Partridge, D. G., Evans, C. M., Freeman, T. M., de Silva, T. I., Angyal, A., Brown, R. L., Carrilero, L., … Montefiori, D. C. (2020). Tracking Changes in SARS-CoV-2 Spike: Evidence that D614G Increases Infectivity of the COVID-19 Virus. *Cell*, *182*(4), 812-827.e19. https://doi.org/https://doi.org/10.1016/j.cell.2020.06.043

Lau, S. K. P., Li, K. S. M., Huang, Y., Shek, C.-T., Tse, H., Wang, M., Choi, G. K. Y., Xu, H., Lam, C. S. F., Guo, R., Chan, K.-H., Zheng, B.-J., Woo, P. C. Y., & Yuen, K.-Y. (2010). Ecoepidemiology and complete genome comparison of different strains of severe acute respiratory syndrome-related Rhinolophus bat coronavirus in China reveal bats as a reservoir for acute, self-limiting infection that allows recombination events. *Journal of Virology*, *84*(6), 2808–2819. https://doi.org/10.1128/JVI.02219-09

Lee, H. J., Shieh, C. K., Gorbalenya, A. E., Koonin, E. V, La Monica, N., Tuler, J., Bagdzhadzhyan, A., & Lai, M. M. (1991). The complete sequence (22 kilobases) of murine coronavirus gene 1 encoding the putative proteases and RNA polymerase. *Virology*, *180*(2), 567–582. https://doi.org/10.1016/0042-6822(91)90071-i

Lei, X., Dong, X., Ma, R., Wang, W., Xiao, X., Tian, Z., Wang, C., Wang, Y., Li, L., Ren, L., Guo, F., Zhao, Z., Zhou, Z., Xiang, Z., & Wang, J. (2020). Activation and evasion of type I interferon responses by SARS-CoV-2. *Nature Communications*, *11*(1), 3810. https://doi.org/10.1038/s41467-020-17665-9

Li, F., Li, W., Farzan, M., & Harrison, S. C. (2005). Structure of SARS coronavirus spike receptor-binding domain complexed with receptor. *Science (New York, N.Y.)*, *309*(5742), 1864–1868. https://doi.org/10.1126/science.1116480

Li, X., Wang, W., Zhao, X., Zai, J., Zhao, Q., Li, Y., & Chaillon, A. (2020). Transmission dynamics and evolutionary history of 2019-nCoV. *Journal of Medical Virology*, *92*(5), 501–511. https://doi.org/https://doi.org/10.1002/jmv.25701

Lu, R., Zhao, X., Li, J., Niu, P., Yang, B., Wu, H., Wang, W., Song, H., Huang, B., Zhu, N., Bi, Y., Ma, X., Zhan, F., Wang, L., Hu, T., Zhou, H., Hu, Z., Zhou, W., Zhao, L., … Tan, W. (2020). Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. *Lancet (London, England)*, *395*(10224), 565–574. https://doi.org/10.1016/S0140-6736(20)30251-8

Lu, W., Zheng, B.-J., Xu, K., Schwarz, W., Du, L., Wong, C. K. L., Chen, J., Duan, S., Deubel, V., & Sun, B. (2006). Severe acute respiratory syndrome-associated coronavirus 3a protein forms an ion channel and modulates virus release. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(33), 12540–12545. https://doi.org/10.1073/pnas.0605402103

Malone, B., Chen, J., Wang, Q., Llewellyn, E., Choi, Y. J., Olinares, P. D. B., Cao, X., Hernandez, C., Eng, E. T., Chait, B. T., Shaw, D. E., Landick, R., Darst, S. A., & Campbell, E. A. (2021). Structural basis for backtracking by the SARS-CoV-2 replication–transcription complex. *Proceedings of the National Academy of Sciences*, *118*(19), e2102516118. https://doi.org/10.1073/pnas.2102516118

Matthew, M., Jessica, B., Anna, D. M., Alex, C., C., W. A., Julia, D. I., Alejandra, T. M., Mary-Jane, N., Chiara, S.-F., Christian, S., R., S. K., Maria, A., Dora, P., Katja, C., Siro, B., Stefano, J., Elisabetta, C., E., B. J., W., T. S., … David, V. (2021). SARS-CoV-2 immune evasion by the B.1.427/B.1.429 variant of concern. *Science*, *373*(6555), 648–654. https://doi.org/10.1126/science.abi7994

McIntosh, K., Dees, J. H., Becker, W. B., Kapikian, A. Z., & Chanock, R. M. (1967). Recovery in tracheal organ cultures of novel viruses from patients with respiratory disease. *Proceedings of the National Academy of Sciences of the United States of America*, *57*(4), 933–940. https://doi.org/10.1073/pnas.57.4.933

Mehta, P., McAuley, D. F., Brown, M., Sanchez, E., Tattersall, R. S., & Manson, J. J. (2020). COVID-19: consider cytokine storm syndromes and immunosuppression. In *Lancet (London, England)* (Vol. 395, Issue 10229, pp. 1033–1034). https://doi.org/10.1016/S0140-6736(20)30628-0

Motozono, C., Toyoda, M., Zahradnik, J., Saito, A., Nasser, H., Tan, T. S., Ngare, I., Kimura, I., Uriu, K., Kosugi, Y., Yue, Y., Shimizu, R., Ito, J., Torii, S., Yonekawa, A., Shimono, N., Nagasaki, Y., Minami, R., Toya, T., … Sato, K. (2021). SARS-CoV-2 spike L452R variant evades cellular immunity and increases infectivity. *Cell Host & Microbe*, *29*(7), 1124-1136.e11. https://doi.org/https://doi.org/10.1016/j.chom.2021.06.006

Myers, E. W., & Miller, W. (1988). Optimal alignments in linear space. *Computer Applications in the Biosciences : CABIOS*, *4*(1), 11—17. https://doi.org/10.1093/bioinformatics/4.1.11

Newman, J. A., Douangamath, A., Yadzani, S., Yosaatmadja, Y., Aimon, A., Brandão-Neto, J., Dunnett, L., Gorrie-stone, T., Skyner, R., Fearon, D., Schapira, M., von Delft, F., & Gileadi, O. (2021). Structure, mechanism and crystallographic fragment screening of the SARS-

CoV-2 NSP13 helicase. *Nature Communications*, *12*(1), 4848. https://doi.org/10.1038/s41467-021-25166-6

Nudler, E. (2012). RNA polymerase backtracking in gene regulation and genome instability. *Cell*, *149*(7), 1438–1445. https://doi.org/10.1016/j.cell.2012.06.003

Padhan, K., Tanwar, C., Hussain, A., Hui, P. Y., Lee, M. Y., Cheung, C. Y., Peiris, J. S. M., & Jameel, S. (2007). Severe acute respiratory syndrome coronavirus Orf3a protein interacts with caveolin. *The Journal of General Virology*, *88*(Pt 11), 3067–3077. https://doi.org/10.1099/vir.0.82856-0

Peiris, J. S. M., Lai, S. T., Poon, L. L. M., Guan, Y., Yam, L. Y. C., Lim, W., Nicholls, J., Yee, W. K. S., Yan, W. W., Cheung, M. T., Cheng, V. C. C., Chan, K. H., Tsang, D. N. C., Yung, R. W. H., Ng, T. K., & Yuen, K. Y. (2003). Coronavirus as a possible cause of severe acute respiratory syndrome. *Lancet (London, England)*, *361*(9366), 1319–1325. https://doi.org/10.1016/s0140-6736(03)13077-2

Peng, Q., Peng, R., Yuan, B., Zhao, J., Wang, M., Wang, X., Wang, Q., Sun, Y., Fan, Z., Qi, J., Gao, G. F., & Shi, Y. (2020). Structural and Biochemical Characterization of the nsp12-nsp7-nsp8 Core Polymerase Complex from SARS-CoV-2. *Cell Reports*, *31*(11), 107774. https://doi.org/https://doi.org/10.1016/j.celrep.2020.107774

Pewe, L., Zhou, H., Netland, J., Tangudu, C., Olivares, H., Shi, L., Look, D., Gallagher, T., & Perlman, S. (2005). A severe acute respiratory syndrome-associated coronavirus-specific protein enhances virulence of an attenuated murine coronavirus. *Journal of Virology*, *79*(17), 11335–11342. https://doi.org/10.1128/JVI.79.17.11335-11342.2005

Pyrc, K., Dijkman, R., Deng, L., Jebbink, M. F., Ross, H. A., Berkhout, B., & van der Hoek, L. (2006). Mosaic structure of human coronavirus NL63, one thousand years of evolution. *Journal of Molecular Biology*, *364*(5), 964–973. https://doi.org/10.1016/j.jmb.2006.09.074

R., W. S., & Sonia, N.-M. (2005). Coronavirus Pathogenesis and the Emerging Pathogen Severe Acute Respiratory Syndrome Coronavirus. *Microbiology and Molecular Biology Reviews*, *69*(4), 635–664. https://doi.org/10.1128/MMBR.69.4.635-664.2005

Raj, V. S., Osterhaus, A. D. M. E., Fouchier, R. A. M., & Haagmans, B. L. (2014). MERS: emergence of a novel human coronavirus. *Current Opinion in Virology*, *5*, 58–62. https://doi.org/10.1016/j.coviro.2014.01.010

Ren, Y., Shu, T., Wu, D., Mu, J., Wang, C., Huang, M., Han, Y., Zhang, X.-Y., Zhou, W., Qiu, Y., & Zhou, X. (2020). The ORF3a protein of SARS-CoV-2 induces apoptosis in cells. *Cellular & Molecular Immunology*, *17*(8), 881–883. https://doi.org/10.1038/s41423-020-0485-9

Ruan, J. and Li, H. (2019) Fast and accurate long-read assembly with wtdbg2. Nat Methods doi:10.1038/s41592-019-0669-3

Ruan, J. and Li, H. (2019) Fast and accurate long-read assembly with wtdbg2. bioRxiv. doi:10.1101/530972

Shu, T., Huang, M., Wu, D., Ren, Y., Zhang, X., Han, Y., Mu, J., Wang, R., Qiu, Y., Zhang, D.-Y., & Zhou, X. (2020). SARS-Coronavirus-2 Nsp13 Possesses NTPase and RNA Helicase Activities That Can Be Inhibited by Bismuth Salts. *Virologica Sinica*, *35*(3), 321–329. https://doi.org/10.1007/s12250-020-00242-1

Siu, K.-L., Yuen, K.-S., Castaño-Rodriguez, C., Ye, Z.-W., Yeung, M.-L., Fung, S.-Y., Yuan, S., Chan, C.-P., Yuen, K.-Y., Enjuanes, L., & Jin, D.-Y. (2019). Severe acute respiratory syndrome coronavirus ORF3a protein activates the NLRP3 inflammasome by promoting TRAF3-dependent ubiquitination of ASC. *FASEB Journal : Official Publication of the Federation of American Societies for Experimental Biology*, *33*(8), 8865–8877. https://doi.org/10.1096/fj.201802418R

Smith, E. C., Blanc, H., Surdel, M. C., Vignuzzi, M., & Denison, M. R. (2013). Coronaviruses lacking exoribonuclease activity are susceptible to lethal mutagenesis: evidence for proofreading and potential therapeutics. *PLoS Pathogens*, *9*(8), e1003565. https://doi.org/10.1371/journal.ppat.1003565

Stadler, K., Masignani, V., Eickmann, M., Becker, S., Abrignani, S., Klenk, H.-D., & Rappuoli, R. (2003). SARS--beginning to understand a new virus. *Nature Reviews. Microbiology*, *1*(3), 209–218. https://doi.org/10.1038/nrmicro775

Su, S., Wong, G., Shi, W., Liu, J., Lai, A. C. K., Zhou, J., Liu, W., Bi, Y., & Gao, G. F. (2016). Epidemiology, Genetic Recombination, and Pathogenesis of Coronaviruses. *Trends in Microbiology*, *24*(6), 490–502. https://doi.org/10.1016/j.tim.2016.03.003

Subissi, L., Posthuma, C. C., Collet, A., Zevenhoven-Dobbe, J. C., Gorbalenya, A. E., Decroly, E., Snijder, E. J., Canard, B., & Imbert, I. (2014). One severe acute respiratory syndrome coronavirus protein complex integrates processive RNA polymerase and exonuclease activities. *Proceedings of the National Academy of Sciences*, *111*(37), E3900 LP-E3909. https://doi.org/10.1073/pnas.1323705111

T., C.-T. C., Lujian, L., R., Y. J., Peter, K., & J., B. M. (2009). Severe Acute Respiratory Syndrome Coronavirus Nonstructural Protein 2 Interacts with a Host Protein Complex Involved in Mitochondrial Biogenesis and Intracellular Signaling. *Journal of Virology*, *83*(19), 10314–10318. https://doi.org/10.1128/JVI.00842-09

Tanner, J. A., Watt, R. M., Chai, Y.-B., Lu, L.-Y., Lin, M. C., Peiris, J. S. M., Poon, L. L. M., Kung, H.-F., & Huang, J.-D. (2003). The Severe Acute Respiratory Syndrome (SARS) Coronavirus NTPase/Helicase Belongs to a Distinct Class of 5′ to 3′ Viral Helicases * . *Journal of Biological Chemistry*, *278*(41), 39578–39582. https://doi.org/10.1074/jbc.C300328200

Tatsuta, T., Model, K., & Langer, T. (2005). Formation of membrane-bound ring complexes by prohibitins in mitochondria. *Molecular Biology of the Cell*, *16*(1), 248–259. https://doi.org/10.1091/mbc.e04-09-0807

TYRRELL, D. A., & BYNOE, M. L. (1965). CULTIVATION OF A NOVEL TYPE OF COMMON-COLD VIRUS IN ORGAN CULTURES. *British Medical Journal*, *1*(5448), 1467–1470. https://doi.org/10.1136/bmj.1.5448.1467

Tyrrel, D. A. J., J. D. Almedia, D. M. Berry, C. H. Cunningham, D. Hamre, M. S. Hofstad, L. Malluci, and K. McIntosh. (1968). Coronavirus. Nature 220:650.

Van Rossum, G., & Drake, F. L. (2009). Python 3 Reference Manual. Scotts Valley, CA: CreateSpace.

Vazquez, C., Swanson, S. E., Negatu, S. G., Dittmar, M., Miller, J., Ramage, H. R., Cherry, S., & Jurado, K. A. (2021). SARS-CoV-2 viral proteins NSP1 and NSP13 inhibit interferon activation through distinct mechanisms. *PLOS ONE*, *16*(6), e0253089. https://doi.org/10.1371/journal.pone.0253089

Viruses, C. S. G. of the I. C. on T. of. (2020). The species Severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2. *Nature Microbiology*, *5*(4), 536–544. https://doi.org/10.1038/s41564-020-0695-z

Wang, S., Nath, N., Adlam, M., & Chellappan, S. (1999). Prohibitin, a potential tumor suppressor, interacts with RB and regulates E2F function. *Oncogene*, *18*(23), 3501–3510. https://doi.org/10.1038/sj.onc.1202684

Washington, N. L., Gangavarapu, K., Zeller, M., Bolze, A., Cirulli, E. T., Schiabor Barrett, K. M., Larsen, B. B., Anderson, C., White, S., Cassens, T., Jacobs, S., Levan, G., Nguyen, J., Ramirez III, J. M., Rivera-Garcia, C., Sandoval, E., Wang, X., Wong, D., Spencer, E., … Andersen, K. G. (2021). Emergence and rapid transmission of SARS-CoV-2 B.1.1.7 in the United States. *Cell*, *184*(10), 2587-2594.e7. https://doi.org/10.1016/j.cell.2021.03.052

Watanabe, Y., Allen, J. D., Wrapp, D., McLellan, J. S., & Crispin, M. (2020). Site-specific glycan analysis of the SARS-CoV-2 spike. *Science (New York, N.Y.)*, *369*(6501), 330–333. https://doi.org/10.1126/science.abb9983

Watanabe, Y., Berndsen, Z. T., Raghwani, J., Seabright, G. E., Allen, J. D., Pybus, O. G., McLellan, J. S., Wilson, I. A., Bowden, T. A., Ward, A. B., & Crispin, M. (2020). Vulnerabilities in coronavirus glycan shields despite extensive glycosylation. *Nature Communications*, *11*(1), 2688. https://doi.org/10.1038/s41467-020-16567-0

Watanabe, Y., Bowden, T. A., Wilson, I. A., & Crispin, M. (2019). Exploitation of glycosylation in enveloped virus pathobiology. *Biochimica et Biophysica Acta. General Subjects*, *1863*(10), 1480–1497. https://doi.org/10.1016/j.bbagen.2019.05.012

World Health Organization. (2020). WHO COVID-19 Dashboard. Retrieved December 12, 2021, from https://covid19.who.int/

Woo, P. C. Y., Lau, S. K. P., Chu, C., Chan, K., Tsoi, H., Huang, Y., Wong, B. H. L., Poon, R. W. S., Cai, J. J., Luk, W., Poon, L. L. M., Wong, S. S. Y., Guan, Y., Peiris, J. S. M., & Yuen, K. (2005). Characterization and complete genome sequence of a novel coronavirus, coronavirus HKU1, from patients with pneumonia. *Journal of Virology*, *79*(2), 884–895. https://doi.org/10.1128/JVI.79.2.884-895.2005

Wuze, R., Xiuxia, Q., Wendong, L., Zhenggang, H., Meng, Y., Peng, Z., Shu-Yi, Z., Lin-Fa, W., Hongkui, D., & Zhengli, S. (2008). Difference in Receptor Usage between Severe Acute Respiratory Syndrome (SARS) Coronavirus and SARS-Like Coronavirus of Bat Origin. *Journal of Virology*, *82*(4), 1899–1907. https://doi.org/10.1128/JVI.01085-07

Xia, S., Lan, Q., Su, S., Wang, X., Xu, W., Liu, Z., Zhu, Y., Wang, Q., Lu, L., & Jiang, S. (2020). The role of furin cleavage site in SARS-CoV-2 spike protein-mediated membrane fusion in the presence or absence of trypsin. *Signal Transduction and Targeted Therapy*, *5*(1), 92. https://doi.org/10.1038/s41392-020-0184-0

Zhai, Y., Sun, F., Li, X., Pang, H., Xu, X., Bartlam, M., & Rao, Z. (2005). Insights into SARS-CoV transcription and replication from the structure of the nsp7–nsp8 hexadecamer. *Nature Structural & Molecular Biology*, *12*(11), 980–986. https://doi.org/10.1038/nsmb999

Zhao, Z., Li, H., Wu, X., Zhong, Y., Zhang, K., Zhang, Y.-P., Boerwinkle, E., & Fu, Y.-X. (2004). Moderate mutation rate in the SARS coronavirus genome and its implications. *BMC Evolutionary Biology*, *4*, 21. https://doi.org/10.1186/1471-2148-4-21

Zhu, N., Zhang, D., Wang, W., Li, X., Yang, B., Song, J., Zhao, X., Huang, B., Shi, W., Lu, R., Niu, P., Zhan, F., Ma, X., Wang, D., Xu, W., Wu, G., Gao, G. F., & Tan, W. (2020). A Novel Coronavirus from Patients with Pneumonia in China, 2019. *New England Journal of Medicine*, *382*(8), 727–733. https://doi.org/10.1056/NEJMoa2001017